1631

RAW SEQUENCE LISTING
PATENT APPLICATION: US/08/939,905B

Input Set : A:\873726u1.sq1
Output Set: N:\CRF3\07112001\H939905B.raw

```
 5 <110> APPLICANT: Gijzen, Mark
 9 <120> TITLE OF INVENTION: Soybean Seed Coat Peroxidase Structural Gene And Regulatory
Region
13 <130> FILE REFERENCE: 75-105
17 <140> CURRENT APPLICATION NUMBER: US 08/939,905B
```

C--> 19 <141> CURRENT FILING DATE: 1997-09-29

```
23 <150> PRIOR APPLICATION NUMBER: US 08/723,414
25 <151> PRIOR FILING DATE: 1996-09-30
29 <160> NUMBER OF SEQ ID NOS: 19
33 <170> SOFTWARE: PatentIn version 2.0
37 <210> SEQ ID NO: 1
39 <211> LENGTH: 1244
41 <212> TYPE: DNA
43 <213> ORGANISM: Glycine max
47 <220> FEATURE:
49 <221> NAME/KEY: CDS
51 <222> LOCATION: (1)..(1056)
55 <220> FEATURE:
57 <221> NAME/KEY: sig_peptide
59 <222> LOCATION: (1)..(78)
63 <400> SEQUENCE: 1
64 atg ggt tcc atg cgt cta tta gta gtg gca ttg ttg tgt gca ttt gct       48
65 Met Gly Ser Met Arg Leu Leu Val Val Ala Leu Leu Cys Ala Phe Ala
66 1               5                  10                  15
68 atg cat gca ggt ttt tca gtc tct tat gct cag ctt act cct acg ttc       96
69 Met His Ala Gly Phe Ser Val Ser Tyr Ala Gln Leu Thr Pro Thr Phe
70             20                  25                  30
72 tac aga gaa aca tgt cca aat ctg ttc cct att gtg ttt gga gta atc     144
73 Tyr Arg Glu Thr Cys Pro Asn Leu Phe Pro Ile Val Phe Gly Val Ile
74         35                  40                  45
76 ttc gat gct tct ttc acc gat ccc cga atc ggg gcc agt ctc atg agg     192
77 Phe Asp Ala Ser Phe Thr Asp Pro Arg Ile Gly Ala Ser Leu Met Arg
78     50                  55                  60
80 ctt cat ttt cat gat tgc ttt gtt caa ggt tgt gat gga tca gtt ttg     240
81 Leu His Phe His Asp Cys Phe Val Gln Gly Cys Asp Gly Ser Val Leu
82 65                  70                  75                  80
84 ctg aac aac act gat aca ata gaa agc gag caa gat gca ctt cca aat     288
85 Leu Asn Asn Thr Asp Thr Ile Glu Ser Glu Gln Asp Ala Leu Pro Asn
86                 85                  90                  95
88 atc aac tca ata aga gga ttg gac gtt gtc aat gac atc aag aca gcg     336
89 Ile Asn Ser Ile Arg Gly Leu Asp Val Val Asn Asp Ile Lys Thr Ala
90             100                 105                 110
92 gtg gaa aat agt tgt cca gac aca gtt tct tgt gct gat att ctt gct     384
93 Val Glu Asn Ser Cys Pro Asp Thr Val Ser Cys Ala Asp Ile Leu Ala
94         115                 120                 125
96 att gca gct gaa ata gct tct gtt ctg gga gga ggt cca gga tgg cca     432
97 Ile Ala Ala Glu Ile Ala Ser Val Leu Gly Gly Gly Pro Gly Trp Pro
98     130                 135                 140
```

```
100 gtt cca tta gga aga agg gac agc tta aca gca aac cga acc ctt gca      480
101 Val Pro Leu Gly Arg Arg Asp Ser Leu Thr Ala Asn Arg Thr Leu Ala
                                         155               160
102 145                  150                                             528
104 aat caa aac ctt cca gca cct ttc ttc aac ctc act caa ctt aaa gct
105 Asn Gln Asn Leu Pro Ala Pro Phe Phe Asn Leu Thr Gln Leu Lys Ala
                                         170               175
106                      165                                             576
108 tcc ttt gct gtt caa ggt ctc aac acc ctt gat tta gtt aca ctc tca
109 Ser Phe Ala Val Gln Gly Leu Asn Thr Leu Asp Leu Val Thr Leu Ser
                                         185               190
110                      185                                             624
112 ggt ggt cat acg ttt gga aga gct cag tgc agt aca ttc ata aac cga
113 Gly Gly His Thr Phe Gly Arg Ala Arg Cys Ser Thr Phe Ile Asn Arg
                         200               205
114            195                                                       672
116 tta tac aac ttt agc aac act gga aac cct gat cca act ctg aac aca
117 Leu Tyr Asn Phe Ser Asn Thr Gly Asn Pro Asp Pro Thr Leu Asn Thr
                                         220
118       210               215                                         720
120 aca tac tta gaa gta ttg cgt gca aga tgc ccc cag aat gca act ggt
121 Thr Tyr Leu Glu Val Leu Arg Ala Arg Cys Pro Gln Asn Ala Thr Gly
                         220               235               240
122 225                                                                 768
124 gat aac ctc acc aat ttg gac ctg agc aca cct gat caa ttt gac aac
125 Asp Asn Leu Thr Asn Leu Asp Leu Ser Thr Pro Asp Gln Phe Asp Asn
                                         250               255
126                      245                                             816
128 aga tac tac tcc aat ctt ctg cag ctc aat ggc tta ctt cag agt gac
129 Arg Tyr Tyr Ser Asn Leu Leu Gln Leu Asn Gly Leu Leu Gln Ser Asp
                                         265               270
130            260                                                       864
132 caa gaa ctt ttc tcc act cct ggt gct gat acc att ccc att gtc aat
133 Gln Glu Leu Phe Ser Thr Pro Gly Ala Asp Thr Ile Pro Ile Val Asn
                         280               285
134            275                                                       912
136 agc ttc agc agt aac cag aat act ttc ttc tcc aac ttt aga gtt tca
137 Ser Phe Ser Ser Asn Gln Asn Thr Phe Phe Ser Asn Phe Arg Val Ser
                                         300
138       290               295                                         960
140 atg ata aaa atg ggt aat att gga gtg ctg act ggg gat gaa gga gaa
141 Met Ile Lys Met Gly Asn Ile Gly Val Leu Thr Gly Asp Glu Gly Glu
                                         315               320
142 305                  310                                            1008
144 att cgc ttg caa aag aat ttt gtg aat gga gac tcg ttt gga tta gct
145 Ile Arg Leu Gln Lys Asn Phe Val Asn Gly Asp Ser Phe Gly Leu Ala
                                         330               335
146                      325                                            1056
148 agt gtg gcg tcc aaa gat gct aaa caa aag ctt gtt gct caa tct aaa
149 Ser Val Ala Ser Lys Asp Ala Lys Gln Lys Leu Val Ala Gln Ser Lys
                                         345               350
150            340                                                      1116
152 taaaccaata attaatgggg atgtgcatgc tagctagcat gtaaaggcaa attaggttgt  1116
154 aaacctcttt gctagctata ttgaaataaa ccaaaggagt agtgtgcatg tcaattcgat  1176
156 tttgccatgt acctcttgga atattatgta ataattattt gaatctcttt aaggtactta  1236
158 attaatca                                                            1244
```

```
170 .221 . LOCATION: (1)..(1532)
171 .210 . FEATURE:
141 .221 . NAME/KEY: sig_peptide
142 .222 . LOCATION: (1533)..(1610)
143 .220 . FEATURE:
144 .221 . NAME/KEY: exon
151 .222 . LOCATION: (1400)..(1751)
145 .220 . FEATURE:
147 .221 . NAME/KEY: exon
148 .222 . LOCATION: (2304)..(2574)
204 .220 . FEATURE:
205 .221 . NAME/KEY: exon
207 .222 . LOCATION: (3405)..(3769)
211 .220 . FEATURE:
213 .221 . NAME/KEY: exon
215 .222 . LOCATION: (4033)..(1511)
219 .220 . FEATURE:
221 .221 . NAME/KEY: Intron
223 .222 . LOCATION: (1752)..(2384)
227 .220 . FEATURE:
229 .221 . NAME/KEY: Intron
231 .222 . LOCATION: (2575)..(3004)
239 .220 . FEATURE:
237 .221 . NAME/KEY: Intron
239 .222 . LOCATION: (3770)..(4032)
243 .220 . FEATURE:
245 .221 . NAME/KEY: CDS
247 .222 . LOCATION: (1513)..(1751)
251 .220 . FEATURE:
252 .221 . NAME/KEY: CDS
255 .222 . LOCATION: (2383)..(2574)
254 .225 . FEATURE:
261 .221 . NAME/KEY: CDS
263 .222 . LOCATION: (3405)..(3769)
267 .220 . FEATURE:
269 .221 . NAME/KEY: CDS
271 .222 . LOCATION: (4033)..(4512)
279 .400 . SEQUENCE:
270 tagataaaaa aatggaaat aatttttctc agatgttgtt tatactgttt ttttaatcag    60
278 aattaaaatt ctctctttaat tatcgacata attttttttg gtgaatatta tcgacataat   120
280 tatttaatat aaattttat tgtaaataga agtgatactt caattttaat attggagaac   180
282 agtacgaaaa atataaaaa ctgttattag aagaaaaaaa tatatggaaa aggttagcta   240
284 catatattat caaaatagt tgttttaatt ggctatataa acccattgt actctttgta   300
286 atttcactt ttcattaa atacatttct acttttaag ttctatattt tctctcaatt   360
288 ttcttcgata aaccaaaaa tttaacatgg tatatcagcg ataccaccca ctttgaaagc   420
290 catgtatgc tagtataagc agcaaaatt tgccctggtt caagcaaagc aagtgttat   480
292 atagatgtta cttttgttga ggaactcatg ccaatggtac tgattgtgaa actgagaaaa   540
294 ataatttgtg gaatttgaat tatatcatt aaatactcct ctcctgacta ccttcgtccc   600
296 tgaaatttgt accatatta tttacaaaa atttgattac aatgcactaa ttaatgaatg   660
```

```
298 tttcttacat tatcatatta tcatatctga cattttgttt ttactttttta taataattat   720
300 tttaaaaagt catacatgca aataattttt taatagttta cagttaaatt tttacagtaa   7
302 aaatgcatga aaattaaact ttattttttcc aagtcatcat ttagtcaaat cccaaaacaa   840
304 tgattatttt ttgcaaatga atgtttattg aacatttaaa tgtagcctaa ttaattctgg   9
306 ttatggtgtc aatgttcaa aacctaatgc aagtcttag caagtacata catagatcta   9
308 attttaaact tatctttacg caagagatat aagattata catctagttt taaacattaa  10.0
310 cttttgtttt tgtgttaaaa aacag'adca tttctttaat tttgtagagt gacgtgctc   10
312 aaccatatta acgaagatt taattgtat tcaagttcat gaacttagta aataagtttt  1200
314 ggtcttcagt tttcaatttt cattacaaca ttatgtaaa atatcaacgt tttctgaaat  1200
316 ttgttgcttg tgtgctccaa ccacatttaa gagattatag aaattaattt tcaagaagat  1200
318 aatgattcct actcttgctg gccctaccat agtacaataa atccactcat aaatcaacaa  1380
320 gtcgtcgtca taggcaattg ggcatcatat cataaacaat acgtacgtga tattatctag  1440
322 tgtctctcag tttacttat gagaaattat ttttctttaa aaaaagttaa ttaataaaaa  1500
324 catttgcgat acgtgagtt acaagaaatc cgccgaattc atctctataa atagaaggat  1558
326 ctatatgaga gataaaatca tattactca aa atg ggt tcc atg cgt ctt tta
                                    Met Gly Ser Met Arg Leu Leu
                                     1               5
328
329 gta gtg gca ttg ttg tgt gca ttt gct atg cat gca ggt ttt tca gtc   1601
331 Val Val Ala Leu Leu Cys Ala Phe Ala Met His Ala Gly Phe Ser Val
             10                 15                 20
332
334 tct tat gct cag ctt act cct acg ttc tac aga gaa aca tgt cca aat   1649
335 Ser Tyr Ala Gln Leu Thr Pro Thr Phe Tyr Arg Glu Thr Cys Pro Asn
336       25                 30                 35
338 ctg ttc cct att gtg ttt gga gta atc ttc gat gct tct ttc acc gat   1697
339 Leu Phe Pro Ile Val Phe Gly Val Ile Phe Asp Ala Ser Phe Thr Asp
340 40                 45                 50                 55
342 ccc cga atc ggg gcc agt ctt atg agg ctt cat ttt cat gat tgc ttt   1745
343 Pro Arg Ile Gly Ala Ser Leu Met Arg Leu His Phe His Asp Cys Phe
344                  60                 65                 70
346 gtt caa gtacgtactt ttttttttcc ttccaaaatg ccctgcatat ttaacaagat   1801
347 Val Gln
350 tgctttgttc acctagaaaa atgtgttttt ttcaacgatc ttacgtacgt ttctttggtt   1861
352 tgaaaaataa atcagagaga gatcaagaaa atagctagaa agaaagcaac gtttttttaa   1921
354 aaggtatttta gtgtagaaa aatattaaaa ctgaagagaa agaaattaaa taagcttttc   1981
356 ttgaatgata tttacatgtc ttattaactt aaagtcacct tttttctttta agttgtgctt   2041
358 gaagaaaaaa gatgctttc agtttagttt tgattaatgc taattatatt tttaattaat   2101
360 taattaatac catatatcta tttaccatat taattattac tatatttcat gatgacaaca   2161
362 gacaagtatt ctaagagagt atcgtagat gattaatttt tttataaaaa aatcttttgc   2221
364 gtgtatagat attctttat aattggtgca gaaacttgta atgctaattg caattaatct   2281
366 tacattgatt aattaatagc tatatcaat atttagttta ggtataggag acaaatcaag   2341
368 tgatctgaac aaattaagtt gttatatttg cattgtgaca g ggt tgt gat gga tca   2397
                                                   Gly Cys Asp Gly Ser
369                                                           75
370
372 gtt ttg ctg aac aac act gat aca ata gaa agc gag caa gat gca ctt   2445
373 Val Leu Leu Asn Asn Thr Asp Thr Ile Glu Ser Glu Gln Asp Ala Leu
374       80                 85                 90
376 cca aat atc aac tca ata aga gga ttg gac gtt gtc aat gac atc aag   2493
377 Pro Asn Ile Asn Ser Ile Arg Gly Leu Asp Val Val Asn Asp Ile Lys
378 95                 100                105                110
```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/939,905B

Input Set : A:\873726u1.sq1
Output Set: N:\CRF3\07112001\H939905B.raw

```
380  aca gcg gtg gaa aat agt tgt cca gac aca gtt tct tgt gct gat att    2541
351  Thr Ala Val Glu Asn Ser Cys Pro Asp Thr Val Ser Cys Ala Asp Ile
386          115              120              125
382                                                                      2594
384  ttt gct att gca gct gaa ata gct tct gtt ctg gtaattaata actcctaatt
385  Leu Ala Ile Ala Ala Glu Ile Ala Ser Val Leu
386          130              135
388  aattcccaac cattaaaaag ttgcatgatt ggattcaaaa ttctatgata ttggggttct   2654
390  gatataaatt tgtgattaaa ttgcagtaaa aaaaattatc atatactttt aataaaaaaa   2714
392  atttatgtaa tttaatttat tattaaaact atttttaaaa ttcaatgcta actctttttt   2774
394  aatcggagca tgtaagctgg cacccaccgt atatcgttgg aagatgcat agaaaccattt  2834
396  aattaatgga tggaatcagt gaaaacattt aattcaaaat actcttattt gtgattagta  2944
398  atcatgttcg ggcaagttac gttgtgtata attaatttga cttaatcaga taaaaaaaca  3014
400  aatggacgca agccggttcg tatagatatc actggcctgt agaatatgtg gttttttcacg 3074
402  tttaaataaa agctagctac tatattatat ttagtctttt tttttcttaa acccatttaa  3134
404  cgtgatttat tgactgtgaa aaatgtttcc acacaagcc ttagaaactc ctcgcaacta   3194
406  acatctccaa aatttgacta tttatttatg aagataattc atctatgatg ttcaactcta  3254
408  ttatatatat gtatcatgc agtattaaga attataatag tcaaatatag agtatatcg    3314
410  ggtaaatgta gttgcatgtg cgacctgttt cgtgtaaaat gcttattcta tatagctttt  3374
412  tttattggaa aataacgatg aactaaaaac gaaagggtat catatatttt gactttatg   3434
414  ttagagagag acatcttaat ttggtcatat gttaaataat taattacaat gcatacacaa  3494
416  atatttatgc catatctaaa aaatgataaa atatcatagg tatactcaac tatatgatat  3554
418  ccccataaca gaaattgtac tttttcttcag gcaatgaact taacattct gtttgctaaa  3610
420  aacaaacatc cacttaaagt ggttcaacat atttatgtaa taatttacag gga gga
                                                                  Gly Gly
421                                                                      3692
424  ggt cca gga tgg cca gtt cca tta gga aga agg gac agt tta aca gca
425  Gly Pro Gly Trp Pro Val Pro Leu Gly Arg Arg Asp Ser Leu Thr Ala    195
426  140              145              150
428  aac cga acc ctt gca aat caa aac ctt cca gca cct ttc ttc aac ctc    3706
429  Asn Arg Thr Leu Ala Asn Gln Asn Leu Pro Ala Pro Phe Phe Asn Leu
430          160              165              170
432  act caa ctt aaa gct tct ttt gct gtt caa ggt ctc aac act ctt gat    3754
435  Thr Gln Leu Lys Ala Ser Phe Ala Val Gln Gly Leu Asn Thr Leu Asp
434          175              180              185
436  tta gtt aca ctc tca ggtatacata atcaattttt tatttgctat tagctagcaa    3809
437  Leu Val Thr Leu Ser
438          190
440  taaaaagtct ctgatacaga catatttaga taaattaatt tctccataaa catttataat  3869
442  aaaattatca atttatgtac ttaaaaatta tggattgaag ctcttttcat ccaactttta  3929
444  ctaaagttaa ggtgcatata atataaaata aactatctct tgttcuttat aaaaagattg  3989
446  aagataagtt aaagtctact tataaatcat taatatagt ata gat ggt cat acg    4044
447                                                        Gly Gly His Thr
448                                                            195
449  ttt gga aga gct cgg tgc agt aca ttc ata aac cga tta tac aac ttc    4092
451  Phe Gly Arg Ala Arg Cys Ser Thr Phe Ile Asn Arg Leu Tyr Asn Phe
452          200              205              210
454  agc aac act gga aac cct gat cca act ctg aac aca aca tac tta gaa    4140
455  Ser Asn Thr Gly Asn Pro Asp Pro Thr Leu Asn Thr Thr Tyr Leu Glu
456          215              220              225
458  gta ttg cgt gca aga tgc ccc cag aat gca act ggg gat aac ctc acc    4158
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/08/939,905B

DATE: 07/11/2001
TIME: 10:00:11

Input Set : A:\873726u1.sql
Output Set: N:\CRF3\07112001\H939905B.raw

L:19 M:271 C: Current Filing Date differs, Replaced Current Filing Date